PARTIALLY COHERENT CONSTELLATIONS FOR
MULTIPLE-ANTENNA SYSTEMS

PRIORITY STATEMENT:

5   This application claims priority to U.S. Provisional Patent Application Serial No. 60/474,476, filed on May 30, 2003, and which is herein incorporated by reference in its entirety.

TECHNICAL FIELD:

10   This invention relates generally to design criteria and construction for signal constellations to be used in systems with imperfect channel state information at the receiver (partially coherent). More particularly, this invention relates to using multiple-antenna vector constellations and design criterion based on the Kullback-Leibler distance between conditional distributions.

15

BACKGROUND:

Wireless communication systems serving stationary and mobile wireless subscribers are currently in wide use and are very popular with consumers. Numerous system layouts and communications protocols have been developed

20   to provide coverage in such wireless communication systems.

The wireless communications channels between the transmit device, or transmission unit, (transmitter) and receive device, or receiver unit, (receiver) are inherently variable. Thus, their quality parameters fluctuate in time. Under

25   favorable conditions, wireless channels exhibit good communication parameters, e.g., large data capacity, high signal quality, high spectral efficiency and throughput. Under these favorable conditions, significant amounts of data can be transmitted via the channel reliably. However, as the channel changes in time, the communication parameters also change. Under altered conditions,

30   former data rates, coding techniques and data formats may no longer be possible. For example, when the channel performance is degraded, the

transmitted data may experience excessive corruption yielding unacceptable communication parameters. For instance, transmitted data can exhibit excessive bit-error rates or packet error rates. The degradation of the channel can be due to a multitude of factors such as general noise in the channel, multi-path fading, loss of line-of-sight path, excessive Co-Channel Interference (CCI) and other factors.

In mobile communications systems, a variety of factors may cause signal degradation and corruption. These include interference from other cellular users within or near a particular cell. Another source of signal degradation is multipath fading, wherein the received amplitude and phase of a signal varies over time. Channel state information at the receiver is usually obtained through a training sequence to offset channel degradation. However, training sequences are an expensive commodity when bandwidth is in demand, because they are typically recurring transmission overhead that does not communicate data that is useful to the end users of the communication system. Limiting or eliminating training sequences necessarily frees bandwidth for other uses.

It has been shown that in a Rayleigh flat-fading environment, the capacity of a multiple-input/multiple-output (MIMO) communication system is increased as compared to a communication system using a single transmit and a single receive antenna. This is because multipath communication offers advantages that can be exploited to increase data rates. Specifically, data capacity increases linearly with the smaller of the number of transmit and receive antennas, provided that the fading coefficients for the multiple sub-channels between transmit and receive antennas are known at the receiver. In a slowly fading channel, where the fading coefficients remain approximately constant for many symbol intervals, the transmitter can send training signals that allow the receiver to accurately estimate the fading coefficients.

In practice, due to the necessarily finite length of the training sequence, there will always be some errors in the channel estimates. In order to maintain a given data rate, more rapidly fading channels would result in shorter training

2

sequences, all other parameters being equal. This is because the data desired by a user would crowd out the training sequence, resulting in even less reliable channel estimates. Employing multiple transmit antennas compounds the above problem by requiring longer training sequences for the same estimation performance, since there are more sub-channels to estimate.

A typical assumption in designing optimal codes and signal constellations is known channel parameters at the receiver. That assumption is especially inappropriate in communication systems with multiple transmit antennas. For fast fading channels where the fading coefficients vary too fast to allow a long training period, or for MIMO systems where very long training sequences are required to accurately train all of the possible channels from the transmitter to the receiver, obtaining an accurate estimate of the channel at the receiver may not always be possible. For the above situations where only a rough estimate of the channel state is available at the receiver, existing signal constellations (e.g. PSK, QAM) and multiple-antenna techniques (e.g., V-BLAST, orthogonal transmit diversity), are no longer optimal because they are designed with the assumption of perfect channel state information at the receiver.

In the presence of channel estimation errors (partially coherent systems), signal constellations which are designed using the statistics of the estimation error are more desirable than those designed for perfect channel state information at the receiver.

Currently, PSK (phase shift key) signal constellations are sometimes used in the case of unreliable channel estimates at the receiver for a single antenna system, because PSK constellations are not sensitive to the errors in the estimates of channel amplitude. However, PSK constellations exhibit poor performance for high rate applications, which require larger signal sets. For a multiple antenna system, conventional constellations (PSK or QAM) are sometimes used in conjunction with some multiple-antenna technique such as V-BLAST or transmit diversity. However, theses approaches assume perfect channel state information at the receiver, which is often an invalid assumption as described

3

above. MIMO communication systems using conventional constellations consequently suffer severe performance degradation in the presence of estimation errors as low as a few percent.

5   What is needed in the art is a new type of signal constellation tailored for the unique challenges of a communication system using multiple transmit and/or multiple receive antennas. Such a constellation system would require short or no training sequences, yet provide acceptable error rates despite imperfect channel state knowledge at the receiver. Ideally, advancement in the art is best

10 served by a technique for designing such a signal constellation to facilitate further refinements.

SUMMARY OF THE PREFERRED EMBODIMENTS:

This invention describes the use of multiple-antenna vector constellations that

15 are optimally designed with the consideration of errors in the channel estimate, to overcome the above-mentioned problems.

This invention describes a design technique for constellations used in a MIMO system when a receiver has only an imperfect estimate of the channel

20 parameters. Such constellations show significant performance improvement over the conventional constellations and MIMO techniques. Also shown is a recursive design technique for the spherical constellations that may be used as subsets of the partially coherent constellations.

25 The constellations of this invention exploit the statistics of the fading and estimation error, to jointly design constellations for MIMOsystems. They use a design metric that is derived for the partially coherent scenarios (as opposed to using the conventional Euclidean design metric which is optimal only when receiver has perfect knowledge of the channel). As a result, they provide

30 substantial performance gains over the conventional techniques in the presence of channel estimation errors as low as a few percent.

4

In accordance with one aspect of the present invention is an n-dimensional space-time signal constellation for use in a multiple-input/multiple-output (MIMO) communication system. The n dimensions are real as opposed to complex. The constellation system may be embodied in or on an electronic, optical, and/or electro-optical storage media. The constellation includes a plurality of constellation points. Each point is disposed within only one and only one of at least two (n-1)-dimensional sub-constellations, wherein $n=2M$ and M is an integer greater than one. Preferably, M is the number of transmit antennas used to send a signal decoded by the n dimensional signal constellation.

The sub-constellations preferably include at least one pair of sub-constellations. Designating the pair of sub-constellations as A and B, for example, the invention preferably provides that each of A and B each defines the same number of points x, and both A and B are disposed symmetrically with respect to each other about an origin of the constellation. In certain instances, the constellation system may include another sub-constellation that defines a maximum number of points y (i.e., more than any other sub-constellation), and is not paired as A and B above. Preferably, each sub-constellation defines a plane that is parallel to all other planes defined by the other sub-constellations.

Additionally, the plurality of points may be dispersed among K subsets, in which case the above characterization of sub-constellations holds for sub-constellations within the same subset. K is an integer greater than one, and preferably the subsets define concentric spheres. Where subsets are present, the nearest distance between points in adjacent subsets is preferably a maximized minimum Kullback-Leibler distance.

In accordance with another aspect of the present invention is a symbol detection method for a receiver of a MIMO communication system. This method includes receiving a multipath signal from M transmit antennas, where M is an integer

greater than one. From the received signal, data samples are obtained. The data samples are then fitted to at least one point of an n-dimensional real signal constellation, wherein $n=2M$. Preferably, the signal constellation is as described above.

In a preferred embodiment of the method above, fitting the data sample to points is accomplished by recursively comparing the data sample to points of sub-constellations within subsets until the data sample is matched to a constellation point. The data sample may be deemed matched once a distance between the constellation point and the data sample diverge as compared to the distance between a previous constellation point and the data sample. In certain embodiments, it may be advantageous to store more than one signal constellation, and selecting which one to use based on the receiver's determination of the number of transmit antennas M that were used to send the communication, or based on a ratio of signal power to noise power of the received signal.

Another aspect of the present invention is a wireless communications system network element such as, for example, a portion of a mobile station, a base station, a receiver symbol detector, and a symbol modulator. The network element includes storage means for storing a digital representation of at least one n-dimensional signal constellation defining a plurality of points. Again, the n dimensions are real dimensions, and $n=2M$, M being an integer greater than one. Each and every on of the plurality of points lies within one and only one of at least two $(n\text{-}1)$-dimensional sub-constellations of points. The sub-constellations may include at least one pair of sub-constellations each defining the same number of x points, the pair being symmetrically disposed about a geometric center or origin of the constellation.

In any of the above, the constellation need not be stored as a physical or geometric body, but rather is preferably stored as an algorithm or mathematical representation of points disposed as above and as further described herein.

BRIEF DESCRIPTION OF THE DRAWINGS:

The foregoing and other aspects of these teachings are made more evident in the following Detailed Description of the Preferred Embodiments, when read in conjunction with the attached Drawing Figures, wherein:

Figure 1 is a block diagram of a wireless communication system with multiple transmit antennas employing a constellation of the present invention.

Figure 2A is a diagram of a three level, three-dimensional spherical constellation according to the preferred embodiment of the present invention.

Figure 2B is a diagram similar to Figure 2A but showing two concentric spheres or subsets of constellation points.

Figure 3 is a graph comparing symbol error rate for several constellations for 4 b/s/Hz, with M=N=2 and $\sigma_E^2 = 0.0$.

Figure 4 is a graph similar to Figure 3 but where $\sigma_E^2 = 0.01$.

Figure 5 is a graph similar to Figure 3 but where $\sigma_E^2 = 0.05$.

Figure 6 is a graph similar to Figure 3 but where $\sigma_E^2 = 0.10$.

Figure 7 is a graph comparing symbol error rate for several constellations for 8 b/s/Hz, with M=N=2 and $\sigma_E^2 = 0.0$.

Figure 8 is a graph similar to Figure 7 but where $\sigma_E^2 = 0.01$.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS:

5   The inventors herein have set forth a design criterion for partially coherent signal constellations that is based on the Kullback-Leibler (KL) distance between conditional distributions. Several single transmit antenna constellations have also been designed by the present inventors, which have been shown to have better performance as compared to conventional constellations. These are

10   described in a paper, herein incorporated by reference, entitled CONSTELLATIONS FOR IMPERFECT CHANNEL STATE INFORMATION AT THE RECEIVER by M.J. Borran, A. Sabharwal, and B. Aazhang, published in the *Proceedings of the 40th Annual Allerton Conference on Communications, Control, and Computing*, Monticello, IL, October 2002.

15

A space-time matrix constellation that is optimally designed with the considers errors in the channel estimate to improve receiver performance in the presence of imperfect channel state information at the receiver. A channel is typically either RF or voice, for transmitting electrical signals between a sending point and

20   a receiving point. Channels are often measured in terms of the amount of spectrum they occupy (bandwidth). Constellations are, for example, graphical representations of signal states for a digital system. Selected phase-amplitude pairs are referred to as constellation points. Constellations of the present invention exploit the statistics of the fading to encode additional information in

25   the amplitudes of the transmit signals (as opposed to the PSK constellations in which all of the constellation points have the same amplitude). This allows for additional points in the constellation (higher rate) with a given peak power. In accordance with the teachings of this invention, and assuming the existence of a given signal-to-noise ratio and estimation variance, a multi-level constellation of

30   desired size is designed using a design criteria based on the Kullback-Leibler (KL) distance between conditional distributions.

When a signal is being received, it has to be demodulated in order for the information therein to be detected. However, a signal transferred over the radio path can be distorted in various ways, thus complicating modulation detection.

5   Signal-impairing phenomena include, *e.g.*, noise and inter-symbol interference (ISI). A signal-distorting phenomenon also arises when a signal on a radio connection is reflected from various obstacles, such as buildings and irregularities in the terrain. In this case, the signal detected at a receiver is the sum of a plurality of propagation paths. Each propagation path is different in

10   length, and signals arrive at the receiver at different points of time, *i.e.* the delay varies. In addition, the movement of a vehicle causes frequency deviations in relation to speed, the deviations being called Doppler frequencies.

One type of modulation that may be used is π/4-DQPSK (π/4-shifted Differential

15   Quaternary Phase Shift Keying modulation). This modulation method comprises eight phase states, but only four phase shifts. Allowed phase shifts (symbols) are +/- π/4 and +/-3π/4. In practice, the π/4-DQPSK constellation varies at intervals of a symbol between two 4-point constellations. Non-idealities of a channel may cause constellation points to shift.

20

It is typical of the radio path that a transmitted signal arrives at a receiver along a plurality of propagation paths, each having a specific time delay. Channel properties also change as a function of time. For example, beams reflected and delayed on the radio path cause inter-symbol interference (ISI). The frequency

25   response, or impulse response, of a channel can be estimated by the use of a discrete-timed filter channel estimator, whose filter tap coefficients model the radio channel. Such a channel estimator is used to describe the state of a radio channel, and refers generally to a mechanism for estimating and maintaining a description of the complex impulse response of a radio channel.

30

Figure 1 shows a communication system 20 that may be used with the present invention including a transmitter 22 and a receiver 24. Typically, each unit 22 and 24 is a transceiver, but are discussed separately herein for simplicity. The

transmitter 22 includes a signal processor 23 coupled to a memory 25 for storing a constellation according to the present invention. The transmitter 22 also includes a number M of transmit antennas 26, M=3 as shown. The receiver includes a number of N receive antennas 28, N=1 as shown. At least one of the integers N or M is greater than one to constitute a MIMO communication system.

The receiver 22 is typically part of a cellular telephone, which has sufficient memory to store signal constellations as look-up tables, algorithms, or combinations thereof, in the telephone handset, or that may retrieve signal constellations that are stored at a transmitter location, such as a base unit location, or, in general, that are stored in any memory that is accessible via a wireless network. The receiver 22 may be used in many cellular telephone applications, one non-limiting example being a cdma2000 cellular telephone system (or evolutions thereof). Upon reception, a signal is received from a transmitter antenna 26 over a channel 30 to a receiver antenna 28 and radio-frequency components, such as a demodulator (not shown), process the signal. Samples are then taken with an A/D converter (not shown) from an intermediate-frequency signal. The samples are applied to a synchronization module, or unit, 32. The synchronization module 32 searches the obtained samples for the training sequence associated with the frame structure and uses it to accurately determine the sampling moment, *i.e.* locations of all symbols in the sample flow.

The synchronization module 32 also controls the radio-frequency components of the receiver so as to maintain a signal arriving at the A/D converter at an optimal level. The synchronization module 32 applies the frame to a channel detector module, or unit, 34.

When information is transferred on a radio channel 30, the signal to be transmitted has to be subjected to modulation. Modulation converts the signal into a form in which it can be transmitted at radio frequency. A modulation method can be considered efficient, for instance, if it allows as much information as possible to be transferred using as narrow a frequency band as possible. Depending on the purpose of use, other features can also be emphasized. Modulation should also cause as little interference as possible to adjacent channels. The channel detector module 34 includes, or is suitably coupled to, a

memory 36.

The detector module 34 is coupled to at least one adaptive channel estimator module or unit 38(a)...(n), where n is any suitable integer number. The channel estimators 38 receive input from the synchronization module 32 via associated interconnectors 40, respectively. Interconnectors 40 are typically wires, or wireless transmission means that are adapted to transmit data. The detector module 32 receives as inputs, outputs from the channel estimators 38, via associated interconnectors 42. Detector module 34 outputs information to estimator modules 38, via associated interconnectors 44. Interconnectors 42 and 44 are similar to interconnectors 40 described herein. Detector module 34 utilizes an algorithm or stored program to demodulate the received signal and compare the demodulated signal to one or more space-time matrix signal constellations, which are typically stored in a memory 36 as a look-up table, an algorithm, or a combination thereof, and preferably stored in the mobile phone handset (also referred to as a mobile station, such as a cellular telephone), but alternatively at a transmitter 22, at a base station, or at any other location that is accessible via a wireless network. A logical channel 30 is formed from the framing unit 46.

An example of the general structure of a receiver 22 has been described to facilitate understanding the present invention. However, the structure of the receiver 22 may change without deviating from the present invention, which is directed to a channel equalizer/detector of a receiver.

It should be noted that the performance gain realized by the present invention becomes substantial as the number of transmit and/or receive antennas 26, 28 increases. A significant improvement in performance is also achieved when the improved signal constellations are used in conjunction with an outer error correcting code. For example, the outer code may be a block or a trellis code designed to encode several signal matrices across time. By designing the outer code based on the Kullback-Leibler (KL) distance criterion, the minimum distance between coded blocks can be further increased, leading to improved

error rate performance.

Design criterion is derived for the very general case of matrix constellations (to be used with MIMO systems over several symbol intervals). Therefore, additional improvements in the performance are obtained when the channel remains constant, or almost constant, for several symbol intervals.

The present invention has application to digital communication in, for example, a Rayleigh flat fading environment using a multiple antenna system. Rayleigh fading is a type of signal fading caused by independent multipath signals having a Rayleigh PDF.

In order to set the parameters of the present invention, it is assumed that the transmitter 22 does not know the channel coefficients, and that the receiver 24 has only an estimate of them with some known estimation variance. Utilizing the Kullback-Leibler (KL) distance between conditional distributions as a performance criterion, a design criterion can be derived based on maximizing the minimum KL distance between constellation points.

A communication system 20 with M transmit antennas 26 and N receive antennas 28 in a block Rayleigh flat fading channel 30 with coherence interval of T symbol periods is modeled using the following complex baseband notation:

$$X = SH + W,\qquad\qquad (1)$$

where S is the $T \times M$ matrix of transmitted signals with power constraint $\sum_{t=1}^{T}\sum_{m=1}^{M}E\{|s_{tm}|^2\} = TP$, where s_{tm}'s are the elements of the signal matrix S, X is the $T \times N$ matrix of received signals, H is the $M \times N$ matrix of fading coefficients, and W is the $T \times N$ matrix of the additive received noise. Elements of H and W are assumed to be statistically independent, identically distributed circular complex Gaussian random variables from the distribution $C\mathcal{N}(0, 1)$. We also assume

12

that $H = \hat{H} + \tilde{H}$, where $\hat{H}$ is known to the receiver but $\tilde{H}$ is not. Furthermore, we assume that $\tilde{H}$ has i.i.d. elements from $C\mathcal{N}(0, \sigma^2)$, and is statistically independent from $\hat{H}$ (this can be obtained, e.g., by using a least minimum mean square error estimator, or LMMSE estimator).

With the above assumptions, the conditional probability density of the received signal can be written as:

$$p(X \mid S, \hat{H}) = E_{\tilde{H}}\{p(X \mid S, \hat{H}, \tilde{H})\} = \frac{\exp\left\{-\operatorname{tr}\left[\left(I_T + \sigma_E^2 SS^H\right)^{-1}\left(X - S\hat{H}\right)\left(X - S\hat{H}\right)^H\right]\right\}}{\pi^{TN} \det^N\left(I_T + \sigma_E^2 SS^H\right)} \quad .(2)$$

Assuming a signal set of size L, $\{S_i\}_{i=1}^L$, and defining $p_i(X) = p(X \mid S_i, \hat{H})$, the Maximum Likelihood (ML) detector for this system will have the following form:

$$\hat{S}_{ML} = S_{\hat{i}_{ML}}, \text{ where } \hat{i}_{ML} = \arg\max_{l \in \{1, \cdots, L\}} p_l(X). \tag{3}$$

Using equation (2) above, the *expected (Kullback-Leibler) KL distance* between signal points S_i and S_j has been derived in the previously incorporated paper as

$$\begin{aligned}
\overline{D}(S_i \parallel S_j) &= N \operatorname{tr}\left\{\left(I_T + \sigma_E^2 S_i S_i^H\right)\left(I_T + \sigma_E^2 S_j S_j^H\right)^{-1}\right\} - NT \\
&\quad - N \ln \det\left\{\left(I_T + \sigma_E^2 S_i S_i^H\right)\left(I_T + \sigma_E^2 S_j S_j^H\right)^{-1}\right\} \\
&\quad + N \ln \det\left\{I_M + \left(1 - \sigma_E^2\right)\left(S_i - S_j\right)^H\left(I_T + \sigma_E^2 S_j S_j^H\right)^{-1}\left(S_i - S_j\right)\right\}
\end{aligned} \tag{4}$$

Adopting the above KL distance as the performance criterion, the signal set design can be formulated as the following optimization problem:

$$\operatorname*{maximize}_{\frac{1}{L}\sum_{l=1}^{L}\|S\|^2 = TP} \min_{i \neq j} \overline{D}(S_i \parallel S_j), \tag{5}$$

where $\|S_i\|^2 = \sum_{t=1}^{T} \sum_{m=1}^{M} |(S_i)_{tm}|^2$ is the total power used to transmit S_i. Since the actual value of N, the number of receiver antennas 28, does not affect the maximization in equation (5), an optimal signal set can be designed assuming $N=1$.

For a fixed spectral efficiency, the constellation size grows exponentially with T. For example, to achieve a spectral efficiency of 4 b/s/Hz with $T = 5$, one needs to design a constellation of $2^{20}=1,048,576$ points. With multiple antennas 26, 28 at the transmitter 22 and receiver 24, even larger spectral efficiency is expected, making the constellation design over multiple symbol intervals even more difficult, and their decoding complexity prohibitively large. For these reasons, the following discussion is limited to constellations for the case of $T=1$, though the concepts and equations may be extended for instances with higher T. One important factor accompanying this assumption is that each transmitted matrix will have a unit rank, and thus will not provide any transmit diversity gain. Therefore, in the absence of channel estimation errors, any transmit diversity scheme is expected to show a better performance. However, as detailed below for the case where channel estimation errors are present, the performance of the described constellations can be significantly better than transmit diversity schemes with comparable computational complexity.

Assuming $T=1$, each S_i will be a complex row vector. The expression for the expected KL distance in equation (4) reduces to:

$$\overline{D}(c_i \| c_j) = \frac{1+\sigma_E^2 \|S_i\|^2}{1+\sigma_E^2 \|S_j\|^2} - 1 - \ln\left(\frac{1+\sigma_E^2 \|S_i\|^2}{1+\sigma_E^2 \|S_j\|^2}\right) + \ln\det\left(I_2 + \frac{1-\sigma_E^2}{1+\sigma_E^2 \|S_j\|^2}(S_i - S_j)^H(S_i - S_j)\right) \quad (6)$$

Using the identity

$$\det(I_M + A_{M \times N} B_{N \times M}) = \det(I_N + B_{N \times M} A_{M \times N}) \quad , \qquad (7)$$

yields

$$\overline{D}(c_i \| c_j) = \frac{1 + \sigma_E^2 \|S_i\|^2}{1 + \sigma_E^2 \|S_j\|^2} - 1 - \ln\left(\frac{1 + \sigma_E^2 \|S_i\|^2}{1 + \sigma_E^2 \|S_j\|^2} \right) + \ln\left(1 + \frac{1 - \sigma_E^2}{1 + \sigma_E^2 \|S_j\|^2} \|S_i - S_j\|^2 \right) . (8)$$

From equation (8), it is apparent that if two constellation points (vectors) have the same norm (i.e., lie on the same M-dimensional complex sphere centered at the origin), the first three terms will cancel out, and the KL distance will be a monotonic function of the Euclidean distance between them. Therefore, considering only constant power constellations (i.e., constellations for which all of the points lie on the same sphere centered at the origin), the design criterion becomes maximizing the minimum Euclidean distance between constellation points, similar to the case of perfect channel state information (CSI) at the receiver. The design problem in this case reduces to the problem of packing points on the surface of an M-dimensional complex sphere (or $2M$-dimensional real sphere).

Conversely, if two points lie on different spheres, the maximized minimum KL distance between them will happen when they lie on a line that passes through the origin, and will be determined by the radii of the two spheres. This means that if one partitions the constellation into subsets of concentric M-dimensional complex spheres, $C_1, \ldots , C_K$, with radii $r_1, \ldots , r_K$, containing $l_1, \ldots , l_K$ points, respectively, and defines the intra-subset and inter-subset distances as

$$D_{intra}(k) = \min_{S_i, S_j \in C_k} \ln\left(1 + \frac{1 - \sigma_E^2}{1 + \sigma_E^2 r_k^2} \cdot 4 r_k^2 \sin^2(\angle \overline{S}_i, \overline{S}_j) \right), \qquad (9)$$

and

$$D_{\mathrm{inter}}(k,k') = \frac{1+\sigma_E^2 r_k^2}{1+\sigma_E^2 r_{k'}^2} - 1 - \ln\left(\frac{1+\sigma_E^2 r_k^2}{1+\sigma_E^2 r_{k'}^2}\right) + \ln\left(1 + \frac{1-\sigma_E^2}{1+\sigma_E^2 r_{k'}^2}\cdot |r_k - r_{k'}|^2\right), \qquad (10)$$

then the maximized minimum KL distance between the constellation points will be greater than or equal to the maximized minimum of the inter-subset and intra-subset KL distances. Signal points along the surface of the same sphere are intra-subset points. An example is any two signal points of Figure 2A. Signal points lying on the surface of different concentric spheres are inter-subset points. An example of inter-subset points is a circular signal point and a triangular signal point of Figure 2B. In equation (9), $\bar{S}_i$ and $\bar{S}_j$ are real vectors constructed by concatenating the real and imaginary parts of S_i and S_j, respectively:

$$\begin{aligned} \bar{S}_i &= [\Re(S_i), \Im(S_i)]^T \\ \bar{S}_j &= [\Re(S_j), \Im(S_j)]^T \end{aligned}, \qquad (11)$$

and $\angle \bar{S}_i, \bar{S}_j$ denotes the angle between the two $2M$-dimensional real vectors.

Therefore, instead of solving the (computationally complex) original optimization in equation (5), the following simplified maximin problem finds a close-to-optimal L-point multilevel constellation of $1 \times M$ vectors with average power P:

$$\underset{1 \le K \le L, \frac{1}{L}\sum_{k=1}^{K} l_k r_k^2 = P, \sum_{k=1}^{K} l_k = L \atop 0 \le r_1 < r_2 < \cdots < r_K}{\text{maximize}} \quad \min\left\{\min_{k=1,\ldots,K} D_{\mathrm{intra}}(k), \min_{k=1,\ldots,K-1} D_{\mathrm{inter}}(k,k+1)\right\}, \qquad (12)$$

where, without loss of generality, it is assumed that $r_1 < r_2 < \ldots < r_K$.

In equation (12), K and $l_1, \ldots, l_K$ are discrete variables, while $r_1, \ldots, r_K$ are

continuous variables. For any fixed value of K and $l_1, \ldots, l_K$ satisfying the specified constraints, equation (12) reduces to a continuous optimization over $r_1, \ldots, r_K$, which can be solved numerically. Even though K is allowed to range from 1 to L in equation (12), a pragmatic solution need not evaluate every possible value for K. Starting from $K = 1$, and increasing the value of K by one each time, we can stop the search once the optimum minimum distance of the solution stops increasing. Moreover, since the intra-subset distance is an increasing function of r_k, the optimum constellation also satisfies the extra constraint $l_1 < l_2 < \ldots < l_K$. This extra constraint can be used to further restrict the domain of the search.

As noted above, the design problem for each subset is equivalent to a packing problem on the surface of an M-dimensional complex ($2M$-dimensional real) sphere. However, since the design and decoding complexities of a maximum packing solution are usually high, a more structured solution for a MIMO communication system is a recursive structure for these spherical constellations which have systematic design and low complexity decoding algorithms.

As an initial matter, let $S_n(L)$ denote the L-point recursively constructed n-dimensional real spherical constellation. Beginning from $n = 2$:

$$S_2(L) = \left\{ \begin{bmatrix} \cos((l-1)2\pi/L) \\ \sin((l-1)2\pi/L) \end{bmatrix} \right\}_{l=1}^{L}. \tag{13}$$

For $n>2$, the constellation is constructed by using a number of $(n-1)$-dimensional recursive sub-constellations as latitudes of the n-dimensional constellation. An example of this procedure for the case of $n=3$ is detailed below.

17

Figure 2A shows a 32-point three-dimensional constellation 50. Figures 2A and 2B are presented to better describe a recursive construction method described herein. The actual signal constellations derived herein as an illustration (whose performance curves are reproduced at Figures 3-8 and described below) are four dimensional (real dimensions) because they reflect a MIMO system with two transmit antennas (M=2), and each transmit antenna requires one complex dimension which is equivalent to two real dimensions. As detailed above, constellations may be designed for any integer M, resulting in a 2M real-dimensioned constellation (an M complex-dimensioned constellation).

The constellation 50 of Figure 2A defines three sub-constellations 52, 54, 56, each disposed in a plane and along a circle circumscribing the surface of the constellation 50. The sub-constellations 52, 54 [$S_2(9)$] nearest the poles of the constellation 50 each define nine discrete points, discrete meaning no points in common. The larger sub-constellation 54 [$S_2(14)$] along the equator of the constellation 50 defines fourteen points. All sub-constellations are discrete respecting all other sub-constellations.

The minimum angle between points in a sub-constellation taken in isolation, such as the points 52a and 52b of the sub-constellation 52, is denoted by β, respecting the geometric center 53 of the sub-constellation 52. However, when a sub-constellation 52 is used as one level of a multi-level constellation 50 as in Figure 2A, the effective angle between the constellation points 32a, 32b is no longer β. With respect to the origin 58 of the constellation 50, the effective minimum angle between points of a sub-constellation 52, 54, 56, α, (i.e., between points on one level) is given by the following equation:

$$\sin\left(\frac{\alpha}{2}\right) = \cos(\theta)\sin\left(\frac{\beta}{2}\right). \qquad (14)$$

18

where θ is the latitude of the sub-constellation as shown in Figure 2A.

Figure 2B depicts a constellation 60 with signal points along the surfaces of K=2 concentric spheres, each sphere being a subset of the entire constellation 60. An inner sphere or inner subset 62 is termed C_1 and each point on its surface also lies within one of three sub-constellations 64, 66, 68 of that inner subset 62. Each of the sub-constellations 64, 66, 68 of the inner subset 62 is depicted as a circle spaced along one of P=3 latitudes or levels at a distance r_1 from the origin 58. Similarly, an outer sphere or outer subset 70 is termed C_2 and each point on its surface also lies within one of three sub-constellations 72, 74, 76 of that outer subset 70. Each of the sub-constellations 72, 74, 76 of the outer subset 70 is depicted as a triangle spaced along one of P=3 latitudes or levels at a distance r_2 from the origin 58. The central sub-constellation 74 of the outer sphere C_2 (outer subset 70) is shown in broken line and not visible through the inner sphere C_1 (inner subset 62) for clarity.

Assuming a spherical constellation (or one of the spherical subsets of a multi spherical constellation) of P levels (sub-constellations), the intra-level and inter-level distances (i.e., the distance between points of the same sub-constellation and the distance between points of disparate sub-constellations, respectively) are defined by the following:

$$d_{intra}(p) = \sin^2(\alpha_p), \quad \text{for } p = 1,\ldots,P, \tag{15}$$

and

$$d_{inter}(p,p') = \sin^2(\theta_p - \theta_{p'}), \quad \text{for } p, p' = 1,\ldots,P, \tag{16}$$

where α_p is the effective minimum intra-level angle of the pth sub-constellation, and θ_p and $\theta_{p'}$ are the latitudes of the sub-constellations p and p', respectively.

Figure 2A depicts graphically the angle α between two adjacent points 52a,

52b of one sub-constellation 52; and the angle θ between two points 52z, 56z,

Similar to the approach above, the optimization problem is simplified by only maximizing the minimum of the intra-level and inter-level distances, instead of maximizing the minimum distance between all pairs of constellation points. The following optimization problem gives that solution:

$$\underset{1\le P\le L,\sum_{p=1}^{P} l_p=L,-\frac{\pi}{2}\le\theta_1<\cdots<\theta_P\le\frac{\pi}{2}}{\text{maximize}} \quad \min\left\{\underset{p=1,\ldots,P}{\min}\,d_{intra}(p),\underset{p=1,\ldots,P-1}{\min}\,d_{inter}(p,p+1)\right\}, \qquad (17)$$

where l_p denotes the number of the points in the pth level or p^{th} sub-constellation.

Similar to the optimization above, here P and l_1, ..., l_p are discrete variables, whereas θ_1, ..., θ_p are continuous variables. For a given choice of P and l_1, ..., l_p satisfying the specified constraints, the optimal values for θ_1, ..., θ_p are found by numerically solving a continuous optimization problem. All of the possible values for P need not be attempted. Starting from $P=1$ and increasing the value of P by one at each time, the search can be stopped once the optimum minimum distance obtained from the above optimizations stops increasing. Also it can be shown that the optimal values for l_1, ..., l_p satisfy the following extra constraint:

$$l_p \ge l_{p-1} \text{ or } l_p \ge l_{p+1}, \quad \text{for } p = 2,\ldots,P-1. \qquad (18)$$

This extra constraint can be used to further restrict the search domain.

For the case of $n>3$, the same procedure as explained above is used, with the difference that instead of S_2 constellations, S_{n-1} constellations are used as the sub-constellations, and the spherical subsets (64, 70 of Figure 2B) of the constellations, if any, are constructed recursively.

20

Simulations were conducted with two different spectral efficiencies of 4 and 8 b/s/Hz. Partially coherent constellations were designed for a 2 x 2 (two-transmit and two-receive antenna) system, and different values of channel estimation variance (1%, 5%, and 10%) were evaluated. Performance of the partially coherent constellations designed in accordance with the above teachings was plotted against other known constellations. Specifically, two independent QPSK or 16QAM constellations were used for the two transmit antennas (resulting in spectral efficiencies of 4 and 8 b/s/Hz, respectively) and plotted as reference curves. These are similar to a V-BLAST scheme (see G. J. Foschini, LAYERED SPACE-TIME ARCHITECTURE FOR WIRELESS COMMUNICATION IN A FADING ENVIRONMENT WHEN USING MULTIPLE ANTENNAS, *Bell Labs Tech. J.,* vol. 1, no. 2, pp. 41-59, 1996), with the difference that the optimal (ML) detector, and not a linear or successive receiver as suggested in the VBLAST scheme, was used for the results reproduced herein in order to have a fair comparison. These results are plotted in Figures 3-6 as "PSK" and in Figures 7 and 8 as QAM.

To compare the proposed schemes with a transmit diversity scheme, the orthogonal transmit diversity scheme of S. M. Alamouti (see A SIMPLE TRANSMIT DIVERSITY TECHNIQUE FOR WIRELESS COMMUNICATIONS, *IEEE Journal on Selected Areas of Communications,* vol. 16, no. 8, pp. 1451-1458, Oct., 1998) was also considered. This diversity scheme has a similar decoding complexity as the one described above, and its performance for a 2 x 2 system with 16QAM and 256QAM constellations (4 and 8 b/s/Hz) was evaluated through simulation. These results are plotted in Figures 3-8 as "Alamouti-QAM" and for brevity are referred to as the Alamouti scheme.

Figures 3-4 shows the symbol error rate curves for the case of 4 b/s/Hz and estimation variances of 0% (Figure 3) and 1 % (Figure 4). In the absence of estimation error (Figure 3), the QPSK and the optimal two-antenna constellations have almost the same performance. Because of higher order of transmit diversity, at high SNR, the Alamouti scheme in this case shows better performance compared to both QPSK and the optimal two-antenna constellations.

With 1% estimation variance (Figure 4), the new constellations according to the present invention start showing consistently better performance for SNR values larger than 15dB. The Alamouti scheme suffers from performance degradation at high SNR, and crosses the curve of the new constellations at around 30dB.

Figures 5-6 shows the symbol error rate curves for the case of 4 b/s/Hz and estimation variances of 5% (Figure 5) and 10% (Figure 6). The performance improvement with the new constellations of the present invention is substantial in these cases. The Alamouti scheme suffers from a severe performance degradation because of channel estimation error, and its performance becomes even worse than the conventional QPSK constellations without any transmit diversity.

Figures 7-8 show the symbol error rate curves for the case of 8 b/s/Hz and estimation variances of 0% (Figure 7) and 1 % (Figure 8). In the absence of estimation error (Figure 7), the 16QAM and the optimal two-antenna constellations have almost the same performance. Though the Alamouti scheme has a larger transmit diversity advantage (indicated by a larger slope in Figure 7), it performs worse than the conventional 16QAM constellations because of smaller coding advantage for SNR values of up to around 23dB.

With 1% estimation variance (Figure 8), the new constellations of the present invention show significant performance improvement compared to the 16QAM constellations. The Alamouti scheme suffers from a severe performance degradation and reaches an error floor of about 3×10^{-1}.

Consider again the receiver 24 of Figure 1, that includes a detector 34. Inputs to the detector 34 include the received signal, a channel estimate from the channel estimators 38a, ...38n, and also possibly the signal to noise ratio (SNR, typically E_b/N_0 for digital signals), some statistic of estimation error, such as variance, σ_E^2, and a signal constellation 50, 60 such as that of Figure 2A or 2B that was previously constructed in accordance with this invention, to include amplitude

22

encoded information based on a fading channel that exploits the statistics of the fading process and the channel estimation error. An output of the detector 34 is a stream of detected symbols. The constellation input may be selected from one of x stored constellations, where x may have a value (typically) in the range of about three to about four representing 3-4 SNR ranges. Each constellation may comprise from a few to several hundred points.

The transmitter 22 of Figure 1 takes a bit stream that is inputted, selects a constellation based on the current SNR from the group of x stored constellations, converts input bits into points from the selected constellation, and modulates the carrier in phase and amplitude, in accordance with the selected constellation point. A symbol corresponding to the inputted bits is transmitted over the channels 30. The current SNR may be made known to the transmitter 22 based on the operation of a power control sub-system, and can be indicated by the receiver 24 through a feedback power control channel. A signal constellation may also be chosen based on the number of transmit antennas M used to send the received signal. The number M may be determined by a header in the received message, by signal processing of characteristics of the received message within the receiver itself, or by other means known in the art. The receiver may select the proper constellation based on pre-programmed logic directing a particular constellation for use with various values of SNR, or by other means known in the art for matching one of several decoding options to a characteristic or quality of the received signal.

At the receiver 24, a symbol is received from the transmitter 22, a constellation is selected from a group of x stored constellations based at least on the current SNR, and the carrier is demodulated, preferably by Maximum Likelihood (ML) demodulation, based on the selected constellation. Hard symbols or soft bits are output, depending on whether the received symbols coded or uncoded.

The constellations used in the present invention may, for example, be implemented as lookup tables in either the transmitter unit 22 and/or the receiver unit 24. Symbol decoding (detection) can be done in two stages of "point in sub-

constellation decoding" and "sub-constellation decoding", similar to trellis coded modulation schemes. That is, given the received signal, the point with the largest likelihood (i.e., the point closest to the received signal) within each sub-constellation is found first (point in sub-constellation decoding). Next, the likelihoods of the best points in different sub-constellations are compared to one another to determine the point having the largest likelihood (sub-constellation decoding). For constellations defining more than one spherical subset, the above two-stage decoding can be iterated for each spherical subset, yielding a three-stage decoding process.

Thus for digital communication in a Rayleigh flat fading environment using a MIMO system, when only partial (imperfect) channel state information is available at the receiver, a partially coherent constellation design based on the Kullback-Leibler distance between distributions has been shown to achieve significant performance gains over the conventional constellations and transmit diversity schemes with comparable complexity.

As used herein, an arcuate surface is a three dimensional curved surface, such as a sphere, an egg, a saddle surface, etc. While non-spherical constellations (or subsets of constellations) may be derived from the teachings herein, retaining a spherical structure allows the KL distance between points of the same spherical subset [put forth in equation (9)] to become a monotonic function of the Euclidean distances between points, rendering constellation design independent from the radius of the spherical subset of the overall constellation. As such, various spherical subsets of points can be designed regardless of radius (using r=1, for example), each spherical subset defining a different number of constellation points. These subsets can be stored, such as in the form of lookup tables, and accessed, assembled, and scaled to reflect a particular SNR for the current transmission or reception. The concentric subsets and sub-constellations of each subset are assembled by optimizing the number of levels (sub-constellations), the number of points in each level, and the radius of each level. So long as the subsets are spherical, the optimization problem is greatly simplified as described herein.

24

While the maximized minimum KL distance between points within a single spherical subset reduces to a maximized minimum Euclidean distance for a spherical constellation, the distance between points of different subsets remains a maximized minimum KL distance, whether or not the subsets are spherical and/or concentric. Equation (10) may be used to optimize the KL distance between such points. Mathematically, no physical structure need be imposed on constellations according to the present invention; the spherical diagrams are for clarity of explanation and to simplify optimization. However, structure greatly simplifies computational complexity, and the multi-level, multi-spherical recursive structure described herein maximally reduces the complexity of constellation optimization.

While described in the context of presently preferred embodiments, those skilled in the art should appreciate that various modifications of and alterations to the foregoing embodiments can be made, and that all such modifications and alterations remain within the scope of this invention. Examples herein are stipulated as illustrative and not exhaustive. As used herein, the term *n*-dimensional real constellation shall refer to a constellation in *n* real dimensions, as opposed to complex dimensions.